



Particulate Air Pollution Exposure and Expression of Viral and Human MicroRNAs in Blood: The Beijing Truck Driver Air Pollution Study

Lifang Hou, Jitendra Barupal, Wei Zhang, Yinan Zheng, Lei Liu, Xiao Zhang, Chang Dou, John P. McCracken, Anaité Díaz, Valeria Motta, Marco Sanchez-Guerra, Katherine Rose Wolf, Pier Alberto Bertazzi, Joel D. Schwartz, Sheng Wang, and Andrea A. Baccarelli

<http://dx.doi.org/10.1289/ehp.1408519>

Received: 7 April 2014

Accepted: 10 June 2015

Advance Publication: 12 June 2015

This article will be available in a 508-conformant form upon final publication. If you require a 508-conformant version before then, please contact ehp508@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.



National Institute of
Environmental Health Sciences

Particulate Air Pollution Exposure and Expression of Viral and Human MicroRNAs in Blood: The Beijing Truck Driver Air Pollution Study

Lifang Hou^{*1,2}, Jitendra Barupal^{*3,4}, Wei Zhang⁵, Yinan Zheng⁶, Lei Liu¹, Xiao Zhang¹, Chang Dou⁷, John P. McCracken³, Anaité Díaz⁸, Valeria Motta⁹, Marco Sanchez-Guerra³, Katherine Rose Wolf¹, Pier Alberto Bertazzi⁹, Joel D. Schwartz³, Sheng Wang¹⁰, and Andrea A. Baccarelli³

¹Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA; ²Robert H. Lurie Comprehensive Cancer Center, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA; ³Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA;

⁴Bioinformatics Infrastructure Facility, University of Rajasthan, Jaipur and Barupal Research Foundation, Jaisalmer, India; ⁵Department of Pediatrics, University of Illinois, Chicago, Illinois, USA; ⁶Institute for Public Health and Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA; ⁷Department of Safety Engineering, China Institute of Industrial Health, Beijing, China; ⁸Center for Health Studies, Universidad del Valle de Guatemala, Guatemala City, Guatemala; ⁹Center of Molecular and Genetic Epidemiology, Department of Environmental and Occupational Health, University of Milan and IRCCS Maggiore Policlinico Hospital, Mangiagalli and Regina Elena Foundation, Milan, Italy;

¹⁰Department of Occupational and Environmental Health, Peking University Health Science Center, Beijing, China. * Contributed equally to the work.

Address correspondence to Lifang Hou, Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, 680 North Lake Shore Drive, Suite 1400, Chicago, Illinois 60611 USA. Telephone: (312) 503-4798. E-mail: l-hou@northwestern.edu

Running title: Particulate air pollution and miRNAs

Acknowledgments: Funding from NIEHS (R21 ES020010 and R21 ES020984-01) and EPA (RD 83479801).

Competing financial interests: The authors declare they have no actual or potential competing financial interests.

Abstract

Background: MicroRNAs (miRNAs) are post-transcriptional gene suppressors and potential mediators of environmental effects. In addition to human miRNAs, viral miRNAs expressed from latent viral sequences are detectable in human cells.

Objective: In a highly exposed population in Beijing, China, we evaluated the associations of particulate air pollution exposure on blood miRNA profiles.

Methods: The Beijing Truck Driver Air Pollution Study (BTDAS) included 60 truck drivers and 60 office workers. We investigated associations of short-term air pollution exposure, using measures of personal PM_{2.5} and (Elemental Carbon) EC, and ambient PM₁₀, with blood NanoString-nCounter miRNA profiles at two exams separated by one to two weeks.

Results: No miRNA was significantly associated with personal PM_{2.5} at a false discovery rate (FDR) of 20%. Short-term ambient PM₁₀ was associated with the expression of 12 miRNAs in office workers only (FDR <20%). Short-term EC was associated with differential expression of 46 human and seven viral miRNAs, the latter including three and four viral miRNAs in office workers and truck drivers, respectively. EC-associated miRNAs differed between office workers and truck drivers with significant effect modification by occupational group. Functional interaction network analysis suggested enriched cellular proliferation/differentiation pathways in truck drivers and pro-inflammation pathways in office workers.

Conclusions: Short-term EC exposure was associated with the expression of human and viral miRNAs that may influence immune responses and other biological pathways. Associations between EC exposure and viral miRNA expression suggest that latent viral miRNAs are potential mediators of air pollution-associated health effects. PM_{2.5}/PM₁₀ exposures showed no consistent relationships with miRNA expression.

Introduction

Air pollution is a major public health concern in the United States and worldwide, accounting for approximately 800,000 deaths annually (WHO 2008). Epidemiological studies have demonstrated that peaks in ambient levels of air particulate matter (PM) pollution, of which traffic is a major source, are rapidly followed, within hours or days after exposure, by increased hospitalization and death, particularly from cardiovascular disease (CVD) (Brook et al. 2010; Wang et al. 2013; Wong et al. 2010) and pulmonary disorders (Laumbach and Kipen 2012). In the presence of stagnating weather, peaks more than 10 times higher than background PM_{2.5} (particulate matter with an aerodynamic diameter <2.5 µm) levels are still frequently recorded in US cities (Bellavia et al. 2013). Inhaled traffic-derived PM up-regulates blood leukocyte pro-inflammatory responses and oxidative stress, two major cardiopulmonary pathways. In blood samples of healthy volunteers collected before, during, and after the 2008 Olympics in Beijing, China, changes in air pollution levels across the different periods were found to be associated with differences in several inflammation biomarkers, such as C-reactive protein (CRP) and fibrinogen (Rich et al. 2012). Blood leukocyte gene expression profiling in 18 PM-exposed welders compared with 10 unexposed controls suggested that that associations between PM and systemic inflammatory responses may be initiated through changes in gene expression (Wang et al. 2005).

MicroRNAs (miRNAs) are small (20–24 nucleotides) non-coding RNAs that regulate gene expression by inducing mRNA cleavage or reducing translation (Baccarelli and Bollati 2009). Inflammation and oxidative stress have been shown to affect miRNA expression in blood leukocytes (Xiao and Rajewsky 2009). Recent *in vitro* studies have shown that diesel exhaust

particles (DEPs), a primary component of PM from vehicular traffic, affect the expression of specific miRNAs (Jardim et al. 2009; Yamamoto et al. 2013). In addition to human miRNAs, miRNAs of viral origin are detectable in human cells (Grundhoff and Sullivan 2011).

Expression of viral miRNAs in human cells may result from active infections or, more often, from the expression of latent retroviral sequences integrated in the host DNA (Grundhoff and Sullivan 2011). MiRNA profiles, particularly viral miRNA profiles, are still understudied in relation to traffic pollution in real-life urban conditions.

In the present study, we investigated two highly exposed groups of 60 truck drivers and 60 office workers in Beijing to examine whether exposure to air pollution, traced through PM_{2.5}, PM₁₀, and traffic elemental carbon (EC) measures, is associated with miRNA profiles in blood leukocytes, measured using the NanoString nCounter miRNA Analysis system, a novel digital technology that can accurately distinguish between highly similar miRNAs with high specificity (Knutsen et al. 2013). Traffic-derived PM and its health implications are a major concern in Beijing due to its high population density and rapid increases in vehicular traffic. The two study groups both had high exposure levels and were selected to sample on different types of exposures: truck drivers are directly exposed to traffic emissions, particularly from diesel exhausts and road dusts; office workers represented the highly exposed urban residential population of Beijing, which has a higher exposure to secondary oxidized traffic particles than truck drivers (Baccarelli et al. 2014; Hou et al. 2014). To identify specific short-term changes in miRNA, we studied each participant on two different examination days one to two weeks apart and assessed each participant's exposure on each examination day using a personal monitor.

Materials and Methods

Study Participants

The Beijing Truck Driver Air Pollution Study (BTDas), conducted between June 15 and July 27, 2008, included two groups with high exposure to air pollution: 60 truck drivers and 60 office workers (Baccarelli et al. 2011). We examined each participant on two separate days separated by a one- to two-week interval. A self-administered questionnaire was used to collect detailed information on demographics, lifestyle, and other exposures. Information on time-varying factors, e.g., smoking, was obtained both for past lifestyle patterns and for consumption on the examination day. Individual written informed consent was provided by the participants and Institutional Review Board approval at all participating institutions was obtained prior to study participant recruitment.

Ambient PM₁₀, Personal PM_{2.5}, and Personal EC Measurements

We used daily averages of ambient PM₁₀ computed from data obtained from 27 monitoring stations distributed representatively across the city to estimate the average PM₁₀ level in Beijing (BMEPB 2008). We measured PM_{2.5} and EC on both examination days using small-weight gravimetric samplers worn by the study participants during the eight hours of work as previously described (Baccarelli et al. 2011). To estimate EC, the blackness of the same filters used to measure PM_{2.5} was assessed using an EEL Model M43D smoke stain reflectometer (Diffusion Systems Ltd., London), applying the standard black-smoke index calculations of the absorption coefficients based on reflectance (ISO 1993). Furthermore, to better characterize the sources of air pollution exposures in truck drivers and office workers, we measured elemental components of PM_{2.5} using a XRF PANalytical Epsilon 5 analyzer (Almelo, Netherlands). For each

element, we calculated enrichment factors (EF) for truck drivers (TD) relative to office workers (OW) as follows: $EF(x)_t = [x_{(TD)}/PM_{2.5(TD)}]/[x_{(OW)}/PM_{2.5(OW)}]$, where $[x_{(TD)}/PM_{2.5(TD)}]$ is the ratio of the mean concentration of element x over the mean $PM_{2.5}$ mass in truck drivers and $[x_{(OW)}/PM_{2.5(OW)}]$ is the ratio of the mean concentration of the same element x over the mean $PM_{2.5}$ mass in office workers. As shown in Supplemental Material, Figure S1, truck drivers were more likely to be exposed to $PM_{2.5}$ enriched in components typically found in vehicle exhausts and road dust, such as calcium (Ca), silicon (Si), chlorine (Cl), etc.

NanoString nCounter Assay for miRNA Profiling

A total of 240 peripheral blood samples from 120 participants was collected in PAXgene Blood RNA Tubes (Qiagen, Valencia, California) at the end of each study day (between 4-6 p.m.). Total RNA was extracted using the PAXgene Blood-RNA Kit Qiagen-763134 (Qiagen, Valencia, California). All samples had optical density ratios $280/260 \geq 1.9$ and $260/230 \geq 1.8$. RIN (RNA Integrity Number) scores showed excellent RNA quality (mean \pm SD: 8.3 ± 0.9). We profiled miRNAs using NanoString nCounter-miRNA expression analysis (NanoString Technologies, Seattle, Washington). This method measures 654 endogenous human-associated miRNAs and 80 viral miRNAs expressed in human cells. Each of the total 21 plates that accommodated the 252 study samples (240 samples plus 12 replicates) included six samples from office workers and six from truck drivers, which were randomized by group labels. Raw data were processed using the NanoStringNorm R package (Waggott et al. 2012). The raw data were \log_2 transformed and normalized using the mean of the six positive controls, which were used to calculate a scaling factor in each column (lane/sample) as suggested by NanoString. The internal positive spike controls were present in each reaction to account for minor differences in hybridization, purification, or binding efficiencies. The data were further background corrected by subtracting

the mean of the six negative controls followed by quantile normalization. Only miRNAs with detectable values in more than 90% of all samples were retained for downstream analyses. Principal components analysis (PCA) was performed on the processed data to identify possible outlier samples. The nCounter miRNA data was also confirmed through cross-platform validation in 20 randomly selected study samples using the TaqMan OpenArray Real-Time PCR Plates (Life Technologies, Carlsbad, California) on the QuantStudio 12K Flex Real-Time PCR System. The average Pearson correlation coefficient was 0.73 (0.63 - 0.79) between the two platforms, thus confirming the robustness of the nCounter platform (see Supplemental Material, Figure S2).

The raw and processed miRNA data have been deposited into the NCBI Gene Expression Omnibus (accession number GSE63087).

Statistical Analysis

Separate mixed-effects regression models were used to estimate associations of within- and between-subject variation in exposure to PM_{2.5}, EC, and PM₁₀ with miRNA expression among all participants combined (pooled analysis), and in truck drivers and office workers separately (SAS 9.1, PROC MIXED, SAS Institute, Inc., Cary NC). We examined within-subject effects representing variation in miRNA expression due to variation in exposure with time, as well as between-subject effects representing variation in miRNA expression due to individual differences in exposures. We fitted models adjusted for variables either not matched or incompletely matched by design between the two groups, i.e., age, sex, body mass index (BMI) (continuous), smoking status, number of cigarettes smoked on the examination day (continuous),

examination day, work hours per week, outdoor temperature (continuous), and dew point (continuous). The mixed-effects model was:

$$\gamma_{ij} = \beta_0 + \beta_{\text{mean}} (\text{Exposure}_{\text{mean}}) + \beta_{\text{diff}} (\text{Exposure}_{\text{diff}}) + \beta_1 X_{1ij} + \dots + \beta_k X_{kij} + \xi_i + e_{ij} \quad [1]$$

where γ_{ij} is the mean expression of a miRNA investigated for the i th subject on the j th examination day ($i=1, \dots, n$; $j=1, 2$); β_0 is the overall intercept; β_{mean} is the regression coefficient for $\text{Exposure}_{\text{mean}}$, i.e., the mean exposure on both work days (between-subject effect); β_{diff} is the regression coefficient for the difference between $\text{Exposure}_{\text{diff}}$, i.e., the difference between exposures on a single work day and mean exposure on both work days (within-subject effect); $\beta_2 \dots \beta_k$ are the regression coefficients for the covariates, which included age, sex, BMI, smoking status, number of cigarettes smoked on the examination day, examination date, work hours per day, and outdoor temperature and dew point on the examination day; ξ_i is the random effect for the subject; and e_{ij} is the residual error term. All tests were two-sided, and miRNA expression changes with a Benjamini-Hochberg false discovery rate (FDR) (Benjamini and Hochberg 1995) $< 20\%$, following previous miRNA expression association studies (Barry et al. 2012; Dorval et al. 2014; Gowrishankar et al. 2014), were considered significant and selected for pathway analysis.

Bioinformatics Analysis

A computationally predicated database, miRBase (Griffiths-Jones et al. 2008), was used as the primary source for gene targets of differentially expressed miRNAs. The Miranda algorithm used in miRBase assigns P -values to individual miRNA-3' UTR target binding site for assessing the statistical significance of predicated complementarity between miRNA and target. Because a gene can have binding sites with different P -values, we used two priori stringent cutoff P -values ($P < 10^{-6}$ and $P < 10^{-8}$) that can provide robust targets for a miRNA. For more

comprehensive and reliable target identification, the annotation was further complemented by two experimental databases, miRTarBase (Hsu et al. 2011) and TarBase (Sethupathy et al. 2006). KEGG (Kyoto Encyclopedia of Genes and Genomes) (Kanehisa et al. 2002) pathway and GO (Gene Ontology) (Gene Ontology 2008) analyses for pathway and biological process enrichment were performed using the combined target genes of all miRNAs with significantly associated with EC (FDR <20%) in each group. To consider enrichment significant, a threshold of Benjamin Hochberg FDR <5% for a KEGG pathway and 1% for GO terms were set with minimum eight genes in a gene set. We created functional interaction (FI) networks using target genes selected with miRBase (cutoff $P < 10^{-8}$), TarBase, and miRTarBase (see Supplemental Material, Figure S3) for each study group. We set the node size according to the node degree to highlight the most-connected genes. We performed pathway enrichment analysis of the genes in the FI networks using the Reactome FI Cytoscape plugin (version 8.2.1), which can be used to construct the networks based on information from both predicated and physical interaction among genes (Wu and Stein 2012). A network-clustering algorithm (Newman 2006) provided by the Reactome FI plugin was used for module identification with a minimum of seven nodes.

Results

Characteristics of Study Participants and Air Pollution Exposure Measurements

The characteristics and personal PM_{2.5}, personal EC, and ambient PM₁₀ exposure levels of the 60 office workers and 60 truck drivers are shown in Table 1. Briefly, truck drivers were older, had higher BMI, higher personal PM_{2.5} and EC levels, smoked more cigarettes during the examination period, and had longer working hours than office workers did.

MiRNA Expression Data

We obtained high correlation of miRNA expression measures among 12 technical replicates using processed data ($r > 0.96$). After background correction and normalization, detectable miRNAs in more than 90% of the 240 total samples were included for further analysis. As a result, 166 miRNAs remained in the analysis, including 159 human and seven viral miRNAs. This number is consistent with the expected detection rate, as indicated in the NanoString nCounter documentation for the miRNA platform (NanoString Technologies 2013). We dropped 12 samples considered obvious outliers due to low RNA yields based on the PCA plot (see Supplemental Material, Figure S4). All 12 samples were from 12 different participants that include eight samples from office workers and four samples from truck drivers. The technical robustness of the NanoString platform was demonstrated by showing that there was no difference in the expression value of these positive controls across the 21 plates ($P > 0.995$, one-way ANOVA) (see Supplemental Material, Figure S5). Negligible plate effect on miRNA expressions was seen in the PCA plot (see Supplemental Material, Figure S6). The mean correlation (Pearson's r) of miRNA expression among paired days was 0.860 for office workers and 0.862 for truck drivers.

MiRNA Signatures Associated with Air Pollution Exposure

Short-term EC average level was associated with differential expression of 46 human miRNAs and all seven viral miRNAs detected in the blood samples (FDR $< 20\%$) using the mixed-effects model. Only seven miRNAs were significantly associated with EC in the pooled analysis. In contrast, when stratified on occupational group, 28 miRNAs were significant for office workers, and 29 for truck drivers, with only 5 miRNAs being common to both groups (hsa-miR-125a-5p,

hsa-miR-1274a, hsa-miR-600, hsa-miR-1283, and hsa-miR-10a) and none being common to all three groups (Figure 1A).

Among the significant miRNAs with FDR <20% in each occupational group, EC was associated with three viral miRNAs (EBV-miR-BART2-5p, EBV-miR-BART6-5p, and KSHV-miR-K12-9) in office workers and four viral miRNAs (EBV-miR-BHRF-1-3, EBV-miR-BHRF-1-1, EBV-BART-9, and HIV1-miR-H1) in truck drivers (see Supplemental Material, Table S1). All seven miRNAs in pooled analysis, and the top 10 significant miRNAs based on *P*-values in each groups are shown in Figure 1B. From the same mixed-effects model, we observed predominant up-regulation of miRNA in truck drivers (28 up-regulated vs. 1 down-regulated) and both up- and down-regulation in office workers (16 up-regulated vs. 12 down-regulated) (see Supplemental Material, Table S1). In a pooled analysis that combined data from both groups, seven miRNAs (all up-regulated) were associated with EC at FDR <20% (Figure 1B and see Supplemental Material, Table S1). For highly significant miRNAs selected at FDR <10% in by group analysis, we observed significant interactions (*P*-Value <0.05) between exposure and occupational groups (see Supplemental Material, Table S2). In a sensitivity analysis excluding current/former smokers (25 participants from office workers and 26 participants from truck drivers), we identified 12 miRNAs in office workers and five miRNAs in truck drivers (FDR <20%) (see Supplemental Material, Table S3). Specifically, 10 miRNAs in office workers (including two viral miRNAs) and two miRNAs in truck drivers (both viral miRNAs) were observed in both analyses using all participants or only non-smokers and, all overlapped miRNAs showed expression change in same direction. These results indicated that the significance of the association of certain miRNAs with EC was not affected by smoking, in spite of the fact that the sample size was limited in the sensitivity analysis. We compared the miRNAs

correlated with EC and other exposure measurements (PM_{2.5} and PM₁₀) and found PM₁₀ had short-term (within-subject) effects on the expression of 12 miRNAs in office workers and one miRNA in the pooled analysis at FDR <20% (see Supplemental Material, Table S4). In truck drivers, we observed no significant miRNAs associated with PM₁₀. Personal exposure to PM_{2.5} was not associated with any miRNAs at FDR <20% (see Supplemental Material, Table S5).

Function and Pathway Analysis of miRNA-Targeted Genes

We identified the target mRNAs of the miRNAs associated with short-term changes in EC levels using multiple databases. We used miRBase to identify the target mRNAs, complemented by miRTarBase and TarBase. Using a threshold of $P < 10^{-6}$, we identified 218 target mRNAs in office workers, 595 target mRNAs in truck drivers, and 219 target mRNAs in the pooled analysis. In contrast, at $P < 10^{-8}$, we identified 168 target mRNAs in office workers, 503 target mRNAs in truck drivers, and 187 target mRNAs in the pooled analysis. Only 8% (61) were common between office workers and truck drivers at $P < 10^{-6}$, and 7% (44) at $P < 10^{-8}$. Ten mRNA targets were common between all three analyses at $P < 10^{-6}$, and nine mRNA targets were common between the three analyses at $P < 10^{-8}$.

We conducted pathway analysis of the gene targets identified in miRBase at either P -value threshold. The pathway analysis showed enrichment of 28 canonical pathways from the KEGG database (14 non-disease pathways and 14 disease pathways) at either miRBase threshold (Figure 2 and see Supplemental Material, Table S6). Notably, two pathways (“colorectal cancer” and “pathways in cancer”) were enriched in the pooled analysis of both groups, as well as in office workers and truck drivers separately. Nine pathways (eight cancer pathways and one cell cycle pathway) were found to be enriched in both office workers and truck drivers. The non-

disease pathways for office workers were related to the immune response, cellular communication and cell growth, while those for truck drivers were largely linked with nucleotide and amino acid metabolic activities (Figure 2 and see Supplemental Material, Table S6). As shown in Figure 2 and Supplemental Material, Figure S7, results from KEGG pathway analysis were robust in terms of the threshold used in miRBase for the identification of mRNA targets. Out of the 18 total pathways identified among office workers, 17 were shared using both thresholds. Out of the 19 pathways identified among truck drivers, 17 were shared using both thresholds (see Supplemental Material, Figure S7).

We also conducted functional analysis based on the GO terms (see Supplemental Material, Figure S8 and Table S7). At any miRbase cutoff, the GO analysis showed that the highest proportion of biological processes was related to biological regulation (three of total five GO terms related to transcription regulation) in the pooled analysis, immune response (more than 35%) in office workers (e.g., Toll-like receptor signaling related pathways, response to drug, double stranded DNA break), and cell growth/death (50%) in truck drivers (e.g., negative regulation of cell cycle, insulin receptor signaling pathways, cell cycle arrest, G1 phase mitotic cell cycle, negative regulation of apoptosis process). Biological processes shared by both groups included pathway related to cell growth/death, biological regulation and metabolism; negative regulation of apoptotic process, negative regulation of cell growth, cell cycle arrest, positive regulation of transcription from RNA polymerase II promoter, negative regulation of transcription from RNA polymerase II promote, and positive regulation of transcription, DNA-dependent. Only one immune or stress response-related GO terms was enriched truck drivers (response to drug). In the combined analysis, no GO term was robust in terms of the miRBase

cutoff. In office workers, 23 of the 31 GO terms were found at both thresholds. In truck drivers, 13 of the 20 GO terms were found at both thresholds (see Supplemental Material, Figure S7).

In contrast, pathways and functional analysis for miRNAs associated with PM₁₀ and PM_{2.5} showed that one KEGG pathway (“pathways in cancer”) and three GO terms (“positive regulation of transcription from RNA polymerase II promoter”, “negative regulation of transcription from RNA polymerase II promoter”, and “negative regulation of cell proliferation”) were enriched in short-term effect of PM₁₀ in office workers (data not shown). For the other analyses on truck drivers and two groups combined in relation to PM₁₀ and PM_{2.5}, no significantly enriched KEGG pathways or GO terms associated with PM₁₀ or PM_{2.5} were found (data not shown).

Pathways and Modules in FI Networks

FI networks (Figure 3) linked 43% of the target mRNAs in office workers and 45% of the target mRNAs in truck drivers by their functional interactions. Pathway enrichment showed 24 pathways in office workers and 22 pathways in truck drivers. The top five pathways were selected based on FDR and visualized in FI networks. Only one of the top five pathways was shared between the two groups. In office workers, the two highest-degree nodes showed gene enrichment in regulation of the P38 alpha and beta pathways and the NOD-like receptor signaling pathway. In truck drivers, the highest-degree node was enriched with genes related to IL-2 and signaling pathways. The network-clustering algorithm showed six modules in office workers and nine modules in truck drivers.

Discussion

In the present study, short-term air pollution exposure from traffic particles, as traced by EC levels, was associated with differential expression of 46 human and seven viral miRNAs, while no consistent association of PM_{2.5} and PM₁₀ exposures with miRNA expressions was observed among three analyses. EC-associated miRNA profiles differed between office workers and truck drivers. Significant miRNAs detected in the stratified analyses by group were masked in the pooled analysis of both groups combined. Among the significant miRNAs (FDR <20%) in each group, EC levels were associated with three viral miRNAs (EBV-miR-BART2-5p, EBV-miR-BART6-5p, and KSHV-miR-K12-9) in office workers and four viral miRNAs (EBV-miR-BHRF-1-3, EBV-miR-BHRF-1-1, EBV-BART-9, and HIV1-miR-H1) in truck drivers. Among human miRNAs, gene function and pathway analyses demonstrated that several of the putative mRNAs targeted by the differentially expressed miRNAs are implicated in inflammation in office workers, and in cellular proliferation- and differentiation-related pathways in truck drivers.

Previous experimental (Jardim et al. 2009) and human studies (Bollati et al. 2010; Motta et al. 2013; Yamamoto et al. 2013) have shown miRNA alterations in response to metal-rich PM, PM_{2.5}, or DEPs. For example, Yamamoto et al. showed altered blood expression of 15 miRNAs in response to DEP exposure (Yamamoto et al. 2013) in asthma patients, including four miRNAs identified in office workers (hsa-miR-142-3p, hsa-miR-21, hsa-miR-125a-5p, and hsa-miR-937) and two miRNAs identified in truck drivers (hsa-miR-125a-5p and hsa-miR-320a) in our study. Our finding on miR-146a in office workers is consistent with data from our previous study of healthy Italian steel workers (Bollati et al. 2010; Motta et al. 2013). However, we also found results that were at variance with previous studies. For instance, in office workers we observed

up-regulation of miR-937 in relation to EC level, while Yamamoto et al. reported down-regulation of miR-937 in relation to DEP exposure (Yamamoto et al. 2013).

The five overlapping miRNAs found in both office workers and truck drivers were all down-regulated. These miRNAs have been reported to participate in cell cycle, proliferation, and apoptosis processes. For example, miR-125a has been shown to target the ErbB pathway in acute myeloid leukemia cells (Ufkin et al. 2014), and miR-10a has been linked to cell proliferation control via targeting of the PI3 kinase pathway (Hu et al. 2014). The combined analysis and the interaction analysis showed effect modifications by group. Truck drivers and office workers have substantially different distributions of changes in personal EC level between the two study days. This difference could have masked the effects in the pooled analysis.

The pathways associated with EC exposures differed between office workers and truck drivers. These discrepancies might originate, at least in part, from miRNA responses specific to different types of exposures or variation in exposure levels, or sample size between the two groups. Our analysis of elemental components showed that, compared with office workers, truck drivers were exposed to PM_{2.5} that was enriched in elements found in road dust. Previous studies suggest that differences in the type of exposure might differentially affect molecular mechanisms; for instance, Hoogendoorn et al. (2012) examined the effect of PM size on RNA and protein expression in normal human derived tracheobronchial epithelial cells and found that different size ranges of PM have variation in number and in the magnitude of expression change in associated genes. Hoogendoorn et al. also detected PM effects on the expression of some of the miRNA target genes observed in our study, such as the proliferation- and carcinogen-related genes *E2F1*, *EGR1* targeted by hsa-miR-21 and hsa-miR-192, and the inflammation-related gene

NFKB1 targeted by miR-146a (Hoogendoorn et al. 2012). These findings suggest that physicochemical properties of air particles may differentially affect miRNA expression.

The most novel finding in the present study was the identification of seven viral miRNAs associated with EC levels, including five Epstein-Barr virus (EBV) miRNAs, one Kaposi's sarcoma-associated herpes virus (KSHV) miRNA, and one human immunodeficiency virus-like (HIV) miRNA. More than 400 viral miRNAs have been discovered from both DNA and RNA viruses, and some of these viral miRNAs are detectable in profiles of RNA extracted from human cells (Grundhoff and Sullivan 2011). The best-characterized viral miRNAs are those coded by EBV, KSHV, and the human cytomegalovirus (HCMV) (Cullen 2006). Latent viruses may regulate host cell miRNAs or encode their own miRNAs using host cell machinery (Kincaid and Sullivan 2012). For instance, EBV can induce expression of miR-155 (Cameron et al. 2008) and miR-146a (Feederle et al. 2011) to mediate B-cell proliferation and inflammation. Virus-encoded miRNAs can target expression of key genes in the cellular pathways of the host (Ramalingam et al. 2012). KSHV-encoded miRNAs can modulate the activity of suppressor proteins involved in leukocyte recruitment (i.e. *ICAM-1*) and up-regulate the antioxidant and angiogenesis gene *HMOX1* (Gallaher et al. 2013). Although studies on the effects of environmental pollutants on miRNAs are accumulating (Bollati et al. 2010; Jardim et al. 2009; Motta et al. 2013), evidence on the effects of air pollution on viral miRNAs has been very limited. The only previous report of effects on a viral miRNA is a non-significant up-regulation of EBV-miR-BHRF-1 (FDR=23%) in the previous study of DEP exposure (Yamamoto et al. 2013). Taken together, these findings may indicate a novel mechanism through which exposure-activated latent viral miRNA may exacerbate cellular effects by suppressing apoptosis, inducing cellular proliferation, and increasing inflammation. Recently, Li et al (2011). observed altered

expression in a viral miRNA coded by HCMV expression in blood cells of subjects with hypertension compared to healthy controls, thus suggesting further potential links to human chronic diseases (Li et al. 2011).

Our study has several notable strengths. We conducted a technical validation of personal PM_{2.5} measures that showed high reproducibility ($r = 0.944$) for our measurements. All participants were evaluated using standardized protocols for blood collection and storage. Blood RNA samples were randomized across plates to limit potential bias from plate effects, and laboratory personnel were blinded to the exposure groups and the study days. We also recognize that our study is subject to a number of limitations. Although we used multivariable models to control potential confounders, we may not be able to exclude residual confounding from unmeasured variables. We used two different thresholds to estimate the robustness of the identifications of target mRNAs in miRBase, which resulted in a number of different results dependent on the threshold used. However, in the KEGG pathway analysis, we found several pathways that were robust to the threshold specifications. The RNA isolated in our study from whole blood is derived from a mixture of cellular and extracellular miRNAs, such as miRNAs from non-leukocytes (i.e., red blood cells) and free miRNAs, which may affect our results. This study is also limited by the lack of validation of key findings in other populations.

In summary, short-term exposure to traffic-derived air particles, as reflected in EC levels, was associated with altered expression of human and viral miRNAs, which may ultimately disturb immune functions and other cellular pathways. These associations were significant only for EC, and profiles differed between office workers and truck drivers. Analyses of personal PM_{2.5} and ambient PM₁₀, as well as pooled analysis of EC in both occupational groups combined, showed

predominantly null results. Our finding on viral miRNAs suggests a possible mechanism through which latent viral miRNAs may exacerbate the adverse cellular effects of traffic air pollution and ultimately contribute to air pollution–associated health effects.

References

- Baccarelli A, Bollati V. 2009. Epigenetics and environmental chemicals. *Curr Opin Pediatr* 21:243-251.
- Baccarelli A, Barretta F, Dou C, Zhang X, McCracken JP, Diaz A, et al. 2011. Effects of particulate air pollution on blood pressure in a highly exposed population in Beijing, China: a repeated-measure study. *Environmental Health* 10:108.
- Baccarelli AA, Zheng Y, Zhang X, Chang D, Liu L, Wolf KR, et al. 2014. Air pollution exposure and lung function in highly exposed subjects in Beijing, China: a repeated-measure study. *Part Fibre Toxicol* 11:51.
- Barry CT, D'Souza M, McCall M, Safadjou S, Ryan C, Kashyap R, et al. 2012. Micro RNA expression profiles as adjunctive data to assess the risk of hepatocellular carcinoma recurrence after liver transplantation. *Am J Transplant* 12:428-437.
- Bellavia A, Urch B, Speck M, Brook RD, Scott JA, Albetti B, et al. 2013. DNA hypomethylation, ambient particulate matter, and increased blood pressure: findings from controlled human exposure experiments. *J Am Heart Assoc* 2:e000212.
- Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological)* 57:289-300.
- BMEPB. 2008. Daily data of ambient PM from 27 monitoring stations distributed representatively across the city from Beijing Municipal Environmental Protection Bureau (BMEPB). Available: <http://www.bjepb.gov.cn/> [accessed 30 May 2015].
- Bollati V, Marinelli B, Apostoli P, Bonzini M, Nordio F, Hoxha M, et al. 2010. Exposure to metal-rich particulate matter modifies the expression of candidate microRNAs in peripheral blood leukocytes. *Environ Health Perspect* 118:763-768.
- Brook RD, Rajagopalan S, Pope CA, 3rd, Brook JR, Bhatnagar A, Diez-Roux AV, et al. 2010. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation* 121:2331-2378.
- Cameron JE, Yin Q, Fewell C, Lacey M, McBride J, Wang X, et al. 2008. Epstein-Barr virus latent membrane protein 1 induces cellular MicroRNA miR-146a, a modulator of lymphocyte signaling pathways. *J Virol* 82:1946-1958.
- Cullen BR. 2006. Viruses and microRNAs. *Nat Genet* 38 Suppl:S25-30.
- Dorval V, Mandemakers W, Jolivet F, Coudert L, Mazroui R, De Strooper B, et al. 2014. Gene and MicroRNA transcriptome analysis of Parkinson's related LRRK2 mouse models. *PloS one* 9:e85510.
- Feederle R, Haar J, Bernhardt K, Linnstaedt SD, Bannert H, Lips H, et al. 2011. The members of an Epstein-Barr virus microRNA cluster cooperate to transform B lymphocytes. *J Virol* 85:9801-9810.
- Gallagher AM, Das S, Xiao Z, Andresson T, Kieffer-Kwon P, Happel C, et al. 2013. Proteomic screening of human targets of viral microRNAs reveals functions associated with immune evasion and angiogenesis. *PLoS Pathog* 9:e1003584.
- Gene Ontology C. 2008. The Gene Ontology project in 2008. *Nucleic Acids Res* 36:D440-444.
- Gowrishankar B, Ibragimova I, Zhou Y, Slifker MJ, Devarajan K, Al-Saleem T, et al. 2014. MicroRNA expression signatures of stage, grade, and progression in clear cell RCC. *Cancer Biol Ther* 15:329-341.

- Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ. 2008. miRBase: tools for microRNA genomics. *Nucleic Acids Res* 36:D154-158.
- Grundhoff A, Sullivan CS. 2011. Virus-encoded microRNAs. *Virology* 411:325-343.
- Hoogendoorn B, Berube K, Gregory C, Jones T, Sexton K, Brennan P, et al. 2012. Gene and protein responses of human lung tissue explants exposed to ambient particulate matter of different sizes. *Inhal Toxicol* 24:966-975.
- Hou L, Zhang X, Zheng Y, Wang S, Dou C, Guo L, et al. 2014. Altered methylation in tandem repeat element and elemental component levels in inhalable air particles. *Environ Mol Mutagen* 55:256-265.
- Hsu SD, Lin FM, Wu WY, Liang C, Huang WC, Chan WL, et al. 2011. miRTarBase: a database curates experimentally validated microRNA-target interactions. *Nucleic Acids Res* 39:D163-169.
- Hu R, Pan W, Fedulov AV, Jester W, Jones MR, Weiss ST, et al. 2014. MicroRNA-10a controls airway smooth muscle cell proliferation via direct targeting of the PI3 kinase pathway. *FASEB J* 28:2347-2357.
- ISO. 1993. ISO 9835, Ambient air – Determination of a black smoke index. Geneva: International Organization for Standardization. Available: <https://www.iso.org/obp/ui/#iso:std:iso:9835:ed-1:v1:en> [accessed 30 May 2015].
- Jardim MJ, Fry RC, Jaspers I, Dailey L, Diaz-Sanchez D. 2009. Disruption of microRNA expression in human airway cells by diesel exhaust particles is linked to tumorigenesis-associated pathways. *Environ Health Perspect* 117:1745-1751.
- Kanehisa M, Goto S, Kawashima S, Nakaya A. 2002. The KEGG databases at GenomeNet. *Nucleic Acids Res* 30:42-46.
- Kincaid RP, Sullivan CS. 2012. Virus-encoded microRNAs: an overview and a look to the future. *PLoS Pathog* 8:e1003018.
- Knutsen E, Fiskaa T, Ursvik A, Jorgensen TE, Perander M, Lund E, et al. 2013. Performance comparison of digital microRNA profiling technologies applied on human breast cancer cell lines. *PloS one* 8:e75813.
- Laumbach RJ, Kipen HM. 2012. Respiratory health effects of air pollution: update on biomass smoke and traffic pollution. *J Allergy Clin Immunol* 129:3-11; quiz 12-13.
- Li S, Zhu J, Zhang W, Chen Y, Zhang K, Popescu LM, et al. 2011. Signature microRNA expression profile of essential hypertension and its novel link to human cytomegalovirus infection. *Circulation* 124:175-184.
- Motta V, Angelici L, Nordio F, Bollati V, Fossati S, Frascati F, et al. 2013. Integrative Analysis of miRNA and inflammatory gene expression after acute particulate matter exposure. *Toxicol Sci* 132:307-316.
- NanoString Technologies. 2013. nCounter® miRNA Analysis in Plasma and Serum Samples. Seattle, Washington: NanoString Technologies, Inc.
- Newman ME. 2006. Modularity and community structure in networks. *Proc Natl Acad Sci U S A* 103:8577-8582.
- Ramalingam D, Kieffer-Kwon P, Ziegelbauer JM. 2012. Emerging themes from EBV and KSHV microRNA targets. *Viruses* 4:1687-1710.
- Rich DQ, Kipen HM, Huang W, Wang G, Wang Y, Zhu P, et al. 2012. Association between changes in air pollution levels during the Beijing Olympics and biomarkers of inflammation and thrombosis in healthy young adults. *JAMA* 307:2068-2078.

- Sethupathy P, Corda B, Hatzigeorgiou AG. 2006. TarBase: A comprehensive database of experimentally supported animal microRNA targets. *RNA* 12:192-197.
- Ufkin ML, Peterson S, Yang X, Driscoll H, Duarte C, Sathyanarayana P. 2014. miR-125a regulates cell cycle, proliferation, and apoptosis by targeting the ErbB pathway in acute myeloid leukemia. *Leuk Res* 38:402-410.
- Waggott D, Chu K, Yin S, Wouters BG, Liu FF, Boutros PC. 2012. NanoStringNorm: an extensible R package for the pre-processing of NanoString mRNA and miRNA data. *Bioinformatics* 28:1546-1548.
- Wang T, Li GX, Sun J, Buys N, Liu HM, Liu MF, et al. 2013. Association between ambient particulate matter and daily cause-specific mortality in Tanggu, Tianjin Binhai New Area, China. *Int J Environ Health Res* 23:205-214.
- Wang Z, Neuburg D, Li C, Su L, Kim JY, Chen JC, et al. 2005. Global gene expression profiling in whole-blood samples from individuals exposed to metal fumes. *Environ Health Perspect* 113:233-241.
- WHO. 2008. World Health Organization (WHO) Fact sheet N°313: Ambient (outdoor) air quality and health. Available: <http://www.who.int/mediacentre/factsheets/fs313/en/> [accessed 30 May 2015].
- Wong CM, Vichit-Vadakan N, Vajanapoom N, Ostro B, Thach TQ, Chau PY, et al. 2010. Part 5. Public health and air pollution in Asia (PAPA): a combined analysis of four studies of air pollution and mortality. *Res Rep Health Eff Inst* 154:377-418.
- Wu G, Stein L. 2012. A network module-based method for identifying cancer prognostic signatures. *Genome Biol* 13:R112.
- Xiao C, Rajewsky K. 2009. MicroRNA control in the immune system: basic principles. *Cell* 136:26-36.
- Yamamoto M, Singh A, Sava F, Pui M, Tebbutt SJ, Carlsten C. 2013. MicroRNA expression in response to controlled exposure to diesel exhaust: attenuation by the antioxidant N-acetylcysteine in a randomized crossover study. *Environ Health Perspect* 121:670-675.

Table 1. Characteristics and exposure levels of the study participants [mean \pm SD or n (%)].

Characteristics	Office Workers (n=60)	Truck Drivers (n=60)	<i>P</i> -value ^a
Personal PM _{2.5} level [$\mu\text{g}/\text{m}^3$]	94.6 \pm 64.9	126.8 \pm 68.8	<0.001
Personal EC level [$\mu\text{g}/\text{m}^3$]	13.1 \pm 4.0	17.3 \pm 6.7	<0.001
Ambient PM ₁₀ level [$\mu\text{g}/\text{m}^3$]	116.7 \pm 50.2	123.5 \pm 50.1	0.29
Age [years]	30.3 \pm 8.0	33.5 \pm 5.7	0.01
Sex			
Male	40 (66.7)	40 (66.7)	
Female	20 (33.3)	20 (33.3)	0.99
BMI [kg/m^2]	22.8 \pm 3.4	24.3 \pm 3.2	0.01
Smoking status			
Never smoker	35 (58.4)	34 (56.7)	
Former	2 (3.3)	2 (3.3)	
Current	23 (38.3)	24 (40.0)	0.99
Cigarettes smoked during examination time ^b	2.6 \pm 5.2	6.4 \pm 9.4	<0.001
Work hours per week	50.6 \pm 11.0	67.3 \pm 14.0	<0.001
Average temperature [$^{\circ}\text{C}$] ^c	25.4 \pm 2.5	25.3 \pm 2.5	0.96
Average dew point [$^{\circ}\text{C}$] ^c	20.6 \pm 2.1	20.6 \pm 2.1	0.93

^a *P*-values were calculated using Student's *t*-test and Fisher's exact test for continuous and categorical variables, respectively.

^b Only current or former smokers.

^c Cumulative of the two study days. Based on 240 total observations (120 study days for office workers and 120 study days for truck drivers). *P*-values were obtained from mixed-effects regression models.

Figures Legends

Figure 1. MiRNAs associated with short-term EC level. (A) Venn diagram showing the number of miRNAs associated with EC level in each analysis group (office workers, truck drivers, and the combined samples). (B) Differentially expressed miRNAs ordered by *P*-value at FDR <20% in the combined samples, and top 10 significant miRNAs in office workers, and in truck drivers. Percent change in miRNA expression is shown on the Y-axis. Up-regulated miRNAs were above the zero % change value and down-regulated miRNAs were below the zero % change value. 95% CIs are represented by error bars.

Figure 2. Pathway analysis. KEGG pathways enriched for gene targets of miRNAs associated with short-term EC level analyzed by Fisher's exact test at FDR <5%. Two threshold cutoffs (10^{-6} and 10^{-8}) for gene selection in the miRBase database are shown on the X-axis. Color of a tile represents enrichment of a pathway; Green=Yes, Red=No. Higher-order KEGG categories are shown on the right side. Higher-order categories that include Human Disease Pathways shown were maroon colored. The KEGG analysis showed enrichment of 28 different pathways (14 non-disease pathways and 14 disease pathways) at either miRBase threshold. Results from pathway analysis were robust to the threshold cutoffs in miRBase for the identification of mRNA targets.

Figure 3. Reactome functional interaction (FI) networks. Reactome FI networks showed interactions of the target mRNAs based on physical or predicated connections. Circles represent nodes and circle size represents node degree (the number of connections one node has). Lines with arrows represent physical interactions between nodes. Each node color represents an enriched pathway, and multi-colored nodes represent genes enriched in multiple pathways. Modules are clusters of nodes with a minimum of seven genes. The top five significant pathways in office workers and truck drivers are listed. The network clustering algorithm showed six modules in office workers and nine modules in truck drivers.

K: KEGG, N: National Cancer Institute, R: Reactome Database.

Figure 1.

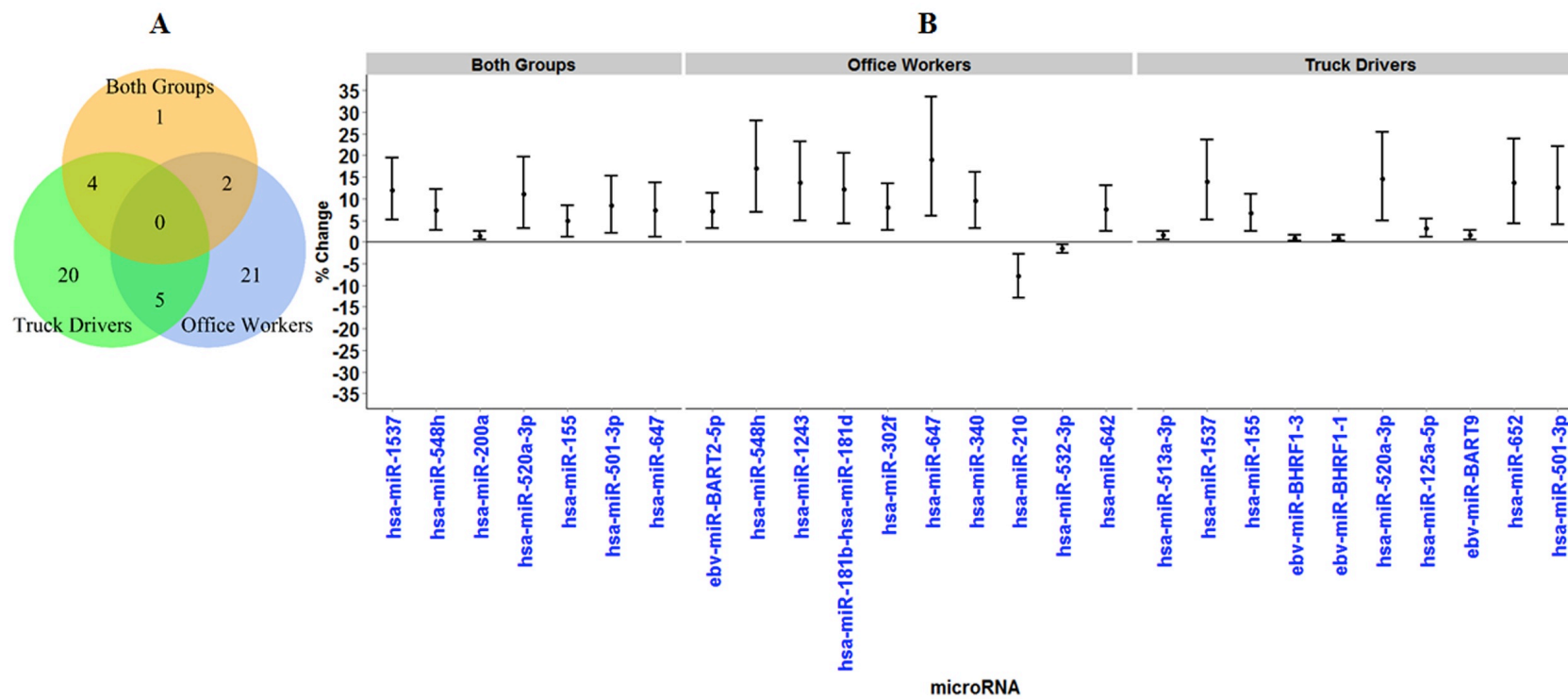


Figure 2.

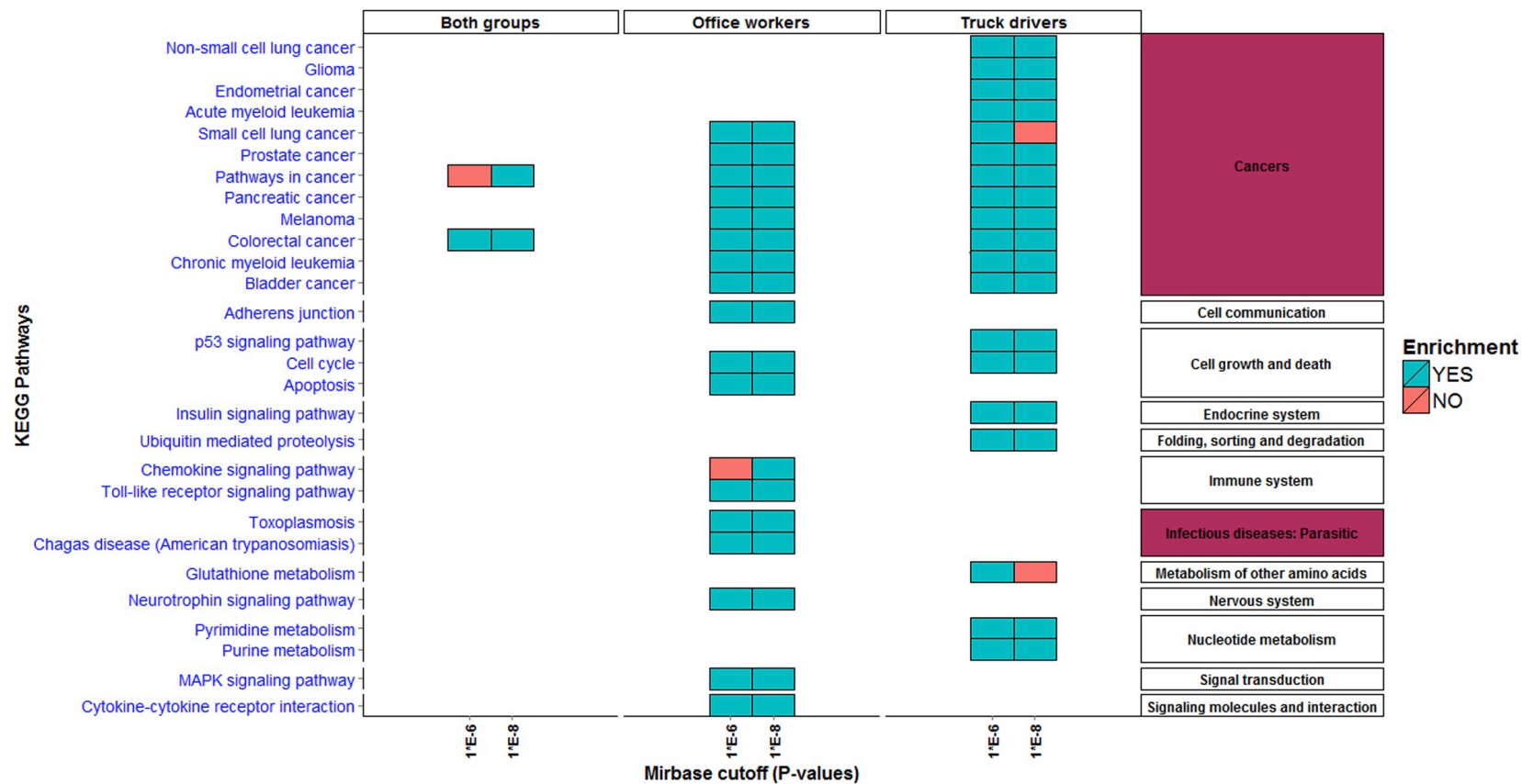


Figure 3.

